

AUGMENTED SPLIT BLOCK EXPERIMENT DESIGN

By

Walter T. Federer

ABSTRACT

A new class of augmented design is presented. These designs are useful for experiments involving screening new treatments over a variety of conditions. The designs are also useful for screening mixtures of materials such as for intercropping experiments. At the later stages of screening material such as genotypes, herbicides, fungicides, etc., it may be desirable to screen over a variety of conditions. The present class of designs allows for this.

Keywords: intercropping; mixing ability; randomized complete block design; incomplete block design; control treatments; new treatments.

BU-1650-M in the Technical Report Series of the Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, New York 14853, April, 2004

INTRODUCTION

A new class of augmented experiment designs is presented herein. This class will allow screening treatments from each of two factors simultaneously, or screening the elements of one factor, e. g. genotypes, under a variety of conditions. This class demonstrates another aspect of the flexibility available to an experimenter as compared to the classes of augmented experiment designs presently available. These designs bear some similarities to the parsimonious designs discussed by Federer and Scully (1993) and Federer (1993), and the class of designs presented by Mejza (1998). A discussion of a member of the class of Mejza designs is discussed along with a statistical analysis. The class of augmented split block experiment designs is presented in the next section. In the following section, the use of these designs in the context of intercropping experiments is discussed. A numerical example is presented to demonstrate the type of analyses that accompany these designs. A few comments are presented in the last section.

Mejza (1998) has presented a class of split block experiment designs wherein a control treatment represents one of the treatments for either or both of the two factors involved. That is, factor A may include a control treatment in every split block experimental unit, sbu, but the sbu treatments change from block to block. Likewise, a control treatment may be included in every sbu for factor B . The following example is

given by Mejza (1998) as an illustration of a member of the class of experiment designs she presented; $A1$ is a control treatment:

<u>Block 1</u>			<u>Block 2</u>			<u>Block 3</u>			<u>Block 4</u>		
	$A1$	$A2$		$A3$	$A1$		$A4$	$A1$		$A1$	$A5$
$B2$	5.37	21.04	$B1$	19.26	1.92	$B1$	18.22	1.15	$B1$	2.10	9.86
$B1$	1.01	14.87	$B2$	26.46	9.97	$B2$	22.23	6.01	$B2$	14.00	27.78

<u>Block 5</u>			<u>Block 6</u>			<u>Block 7</u>			<u>Block 8</u>		
	$A2$	$A1$		$A3$	$A1$		$A1$	$A4$		$A5$	$A1$
$B2$	27.29	5.97	$B1$	16.48	2.20	$B1$	1.18	17.98	$B2$	24.50	6.14
$B1$	24.30	1.29	$B2$	26.22	14.66	$B2$	8.09	25.15	$B1$	14.24	1.10

The control $A1$ is replicated eight times whereas $A2$ to $A5$ are only replicated twice. $B1$ and $B2$ treatments are in a standard split block arrangement of eight blocks. A Type I analysis of variance for this data set is given below:

<u>Source of variation</u>	<u>Degrees of freedom</u>	<u>Sum of squares</u>	<u>Mean square</u>
Total	32	8,216.7179	
Correction for mean	1	5,461.1700	
Block	7	72.0031	10.2862
A	4	2,042.2716	510.5679
$A \times \text{block}$	4	47.8008	11.9502
B	1	478.3324	478.3324
$B \times \text{block}$	7	95.5070	13.6439
$A \times B$	4	17.4306	4.3577
$A \times B \times \text{block}$	4	2.2024	0.5506

A Type III analysis of variance for the above data set is:

<u>Source of variation</u>	<u>Degrees of freedom</u>	<u>Sum of squares</u>	<u>Mean square</u>
Block	7	80.0863	11.4409
A	4	2,042.2716	510.5679
$A \times \text{block}$	4	47.8008	11.9502
B	1	376.7054	376.7054
$B \times \text{block}$	7	56.6374	8.0911
$A \times B$	4	17.4306	4.3577
$A \times B \times \text{block}$	4	2.2024	0.5506

The $A \times \text{block}$ mean square is used as error A , the $B \times \text{block}$ mean square is the error for factor B , and the $A \times B \times \text{block}$ mean square is the error term for the $A \times B$ interaction. A SAS model to obtain the above analyses is:

Yield = block A $A \times \text{block}$ B $B \times \text{block}$ $A \times B$ $A \times B \times \text{block}$;

Using only blocks 1 to 4 or 5 to 8 would not result in a solution for all effects. However an arrangement similar to the following does result in solutions for effects:

Block 1

	<i>A1</i>	<i>A2</i>	<i>A3</i>
<i>B2</i>	5.37	21.04	27.29
<i>B1</i>	1.01	14.87	24.30

Block 2

	<i>A3</i>	<i>A1</i>	<i>A4</i>
<i>B1</i>	19.26	1.92	5.97
<i>B2</i>	26.46	9.97	1.29

Block 3

	<i>A4</i>	<i>A1</i>	<i>A5</i>
<i>B1</i>	18.22	1.15	2.10
<i>B2</i>	22.23	6.01	14.00

Block 4

	<i>A1</i>	<i>A5</i>	<i>A2</i>
<i>B1</i>	2.10	9.86	24.50
<i>B2</i>	14.00	27.78	14.24

A Type III analysis of variance for the above data set is:

Source of variation	Degrees of freedom	Sum of squares	Mean square
Block	3	164.5603	54.8534
<i>A</i>	4	1,159.4945	289.8736
<i>A</i> × block	4	268.1266	67.0316
<i>B</i>	1	147.3927	147.3927
<i>B</i> × block	3	3.1034	1.0345
<i>A</i> × <i>B</i>	4	177.2756	44.3189
<i>A</i> × <i>B</i> × block	4	114.9127	28.7282

The same error terms and SAS model apply here as for the previous data set.

AUGMENTED SPLIT BLOCK EXPERIMENT DESIGNS

The designs of Mejza (1998) have some similarities to augmented experiment designs (See Federer, 1993, 2002, and related references therein). Usually the "new" treatments are only included once in an experiment. In Mejza's (1998) design, *A2* to *A5* were included twice while the control, *A1*, was replicated eight times. One could use the above ideas to construct a class of *augmented split block experiment designs*. For example, suppose that an experimenter used $ac = 4$ standard or control genotypes and $cn = 64$ new genotypes as the *A* factor and two levels of fertilizers with two types of soil preparation as the *B* factor. It may be desirable to replicate the new genotypes only once in an experiment. It may be that this experiment is to be conducted at several sites or locations that have certain characteristics. It is desired to test or screen the new treatments on the $b = 4$ *B* treatments. Further suppose that the designs for the *A* and *B* factors were an augmented randomized complete block design and a randomized complete block experiment design, respectively. Given that eight complete blocks, replicates, are to be used and that eight of the 64 new *A* treatments are to be included in each of the eight blocks as the augmented treatments, an analysis of variance table with a partitioning of the degrees of freedom is:

Source of variation	Degrees of freedom	
	Example	General
Total	320	$rb(ac - an / r)$
Correction for the mean	1	1
Replicate = R	7	$r - 1$
A treatments	67	$ac + an - 1 = a - 1$
Control	3	$ac - 1$
Control versus new	1	1
New	63	$an - 1$
$A \times R = \text{error } A$	21	$(ac - 1)(r - 1)$
B	3	$b - 1$
Soil	1	$s - 1$
Fertilizer	1	$f - 1$
Soil \times fertilizer	1	$(s - 1)(f - 1)$
$B \times R$	21	$(b - 1)(r - 1)$
$A \times B$	201	$(a - 1)(b - 1)$
Control $\times B$	9	$(ac - 1)(b - 1)$
Control vs. new $\times B$	3	$(1)(b - 1)$
New $\times B$	189	$(an - 1)(b - 1)$
$A \times B \times R$	63	$(ac - 1)(b - 1)(r - 1)$

Suppose that the conditions of the above example hold but that an incomplete block design with r incomplete blocks of size $k = 2$ is used for the $ac = 4$ A control treatments. A plan before randomization for this could be:

Block 1		Block 2		Block 3		Block 4	
$A1$	$A2$	$A2$	$A3$	$A3$	$A4$	$A4$	$A1$
$B1$		$B1$		$B1$		$B1$	
$B2$		$B2$		$B2$		$B2$	
$B3$		$B3$		$B3$		$B3$	
$B4$		$B4$		$B4$		$B4$	

In the above, factor A experimental units, sixteen of the 64 new treatments would be included in each of the four blocks. A set of sixteen new and two controls would be randomly allotted to the eighteen experimental units for factor A in each of the four blocks. The B treatments would run across these eighteen A treatments in each of the $r = 4$ blocks. The B treatments are in a randomized complete block design arrangement.

A partitioning of the degrees of freedom for the above designed experiment would be:

Source of variation	Degrees of freedom	
	Example	General
Total	288	$anb + rkb$
Correction for mean	1	1
Block	3	$r - 1$
A treatments	67	$ac + an - 1 = a - 1$
Control or standard	3	$ac - 1$
Control versus new	1	1
New	63	$an - 1$
$A \times \text{block} = \text{error for } A$	1	$rk - ac - r + 1$
B	3	$b - 1$
$B \times \text{block} = \text{error for } B$	9	$(b - 1)(r - 1)$
$A \times B$	201	$(a - 1)(b - 1)$
$B \times \text{control}$	9	$(ac - 1)(b - 1)$
$B \times \text{control versus new}$	3	$b - 1$
$B \times \text{new}$	189	$(b - 1)(an - 1)$
$A \times B \times \text{block} = \text{error for } A \times B$	3	$(rk - ac - r + 1)(b - 1)$

Though the above design is connected, there are insufficient degrees of freedom associated with the error terms. Hence, additional blocks will be required. The three interaction with block mean squares would form the error terms for this augmented split block experiment design. If more new treatments are to be tested, the number of blocks could be increased or more new treatments could be included in the blocks of the factor A experimental units.

As a further variation of an augmented split block experiment design, consider the following arrangement where both factors are designed as augmented randomized complete blocks:

block 1								block 2										block 3												
factor A								factor A										factor A												
factor B	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	factor B	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>8</u>	<u>9</u>	<u>10</u>	factor B	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>11</u>	<u>12</u>	<u>13</u>							
1								1								1														
2								2								2														
3								3								3														
4								6								8														
5								7								9														

Here 1, 2, 3, and 4 are the control treatments for factor A and 5 to 13 are the augmented or new treatments included once. For the B factor, 1, 2, and 3 are the control treatments for this factor and 4 to 9 are the new treatments. An analysis of variance partitioning of the 105 degrees of freedom is:

Source of variation	Degrees of freedom	
	Example	General
Total	3(5)(7) = 105	$r(ac + an/r)(bc + bn/r)$
Correction for the mean	1	1
Block = R	2	$r - 1$
A	12	$ac + an - 1$
Control = AC	3	$ac - 1$
New = AN	8	$an - 1$
AC versus AN	1	1
$A \times R$ = error A	2(3) = 6	$(ac - 1)(r - 1)$
B	8	$bc + bn - 1$
Control = BC	2	$bc - 1$
New = BN	5	$bn - 1$
BC versus $BN = BCN$	1	1
$B \times R$ = error B	2(2) = 4	$(bc - 1)(r - 1)$
$A \times B$	60	$ac(bc + bn - 1) + (an - 1)(bc - 1) + bn(an/r - 1)$
$AC \times BC$	3(2) = 6	$(ac - 1)(bc - 1)$
$AC \times BN$	3(5) = 15	$(ac - 1)(bn - 1)$
$AC \times BC$ vs. AN	3(1) = 3	$(ac - 1)(1)$
$AN \times BC$	8(2) = 16	$(an - 1)(bc - 1)$
$AN \times BN$ / block	2(1)(3) = 6	$r(an / r - 1)(bn / r - 1)$
$AN \times BC$ vs. BN / block	2(1)(3) = 6	$r(1)(an / r - 1)$
AC vs. $AN \times BC$	1(2) = 2	1($bc - 1$)
AC vs. $AN \times BN$ / block	1(1)(3) = 3	1($bn / r - 1$)(r)
AC vs. $AN \times BC$ vs. BN / block	1(1)(3) = 3	r
$A \times B \times R$ = error AB	3(2)(2) = 12	$(ac - 1)(bc - 1)(r - 1)$

Note that not all interactions of new treatments are obtained. The interactions of new treatments in the same block are obtainable as there are two new B treatments and three new A treatments in any block. For some of the interactions, the contrasts vary from block to block and will be different. This is why some of the degrees of freedom were partitioned by block.

If for example, the B treatments were to go across all three blocks of the previous example and the A treatments stayed the same, then there would be a total $9(4 + 3)(3) = 189$ experimental units. The partitioning of the degrees of freedom for this situation is:

Source of variation	Degrees of freedom
Total	189
Correction for the mean	1
Block = R	2
Factor A = error A	12
Control = AC	3
New = AN	8
AC versus AN	1

$A \times R$	$2(3) = 6$
Factor B	8
BC	2
BN	5
BC vs. BN	1
$B \times R$ (not an error term)	$8(2) = 16$
$B \times A$	$8(12) = 96$
$B \times AC$	$8(3) = 24$
$B \times AN$	$8(8) = 64$
$B \times AC$ vs. AN	$8(1) = 8$
$A \times B \times R = AC \times B \times R = \text{error } AB$	$3(8)(2) = 48$

For this arrangement, all interaction terms are available.

AUGMENTED SPLIT BLOCKS FOR INTERCROPPING EXPERIMENTS

Consider the next to the last arrangement in the previous section. Here the treatments 1, 2, 3, and 4 of factor A could be four standard maize cultivars and A treatments 5 to 13 could be nine promising new genotypes of maize. Factor B treatments 1, 2, and 3 could be three standard bean cultivars and treatments 4 to 9 could be six promising new bean genotypes. The goal could be to determine how these mixtures of maize and beans perform as an intercrop. The maize could be planted and then perpendicular to the maize rows, the rows of beans could be planted. Here not all possible mixtures result. If this is desired, then the last example of the previous section could be used.

In intercropping and other mixture experiments, responses may be available for the mixture or for individual components of the mixture. If the latter responses are available, analyses may be conducted for each of the components. For example in intercropping experiments with maize and beans, responses for maize and for beans may be obtained. These responses may also be combined and an analysis is conducted on the combined responses. When responses are available for each crop in the mixture, it is possible to estimate the mixing effects for each component of the mixture (See Federer, 1993, 1999).

Suppose that an experimenter desired to screen one set of new genotypes, say beans, and another set of new genotypes, say maize, to determine their suitability for intercropping systems. The treatments for the A factor could be the controls and new genotypes of maize and the treatments for factor B could be the controls and new genotypes for beans. Further suppose that an augmented randomized complete block experiment design was used for both the A and B factors with $r = 8$ blocks, say. Let the number of new maize genotypes be 100 and the number of new bean genotypes be 96. Twelve new bean genotypes with the four bean controls are randomly allotted to the 16 B experimental units in each of the $r = 8$ blocks, replicates. Given that five maize genotypes are the controls, there will be $5 + 12 = 17$ A experimental units in four of the blocks and $5 + 13 = 18$ experimental units in the other four blocks to accommodate the 100 new maize genotypes. These would be randomly allotted in each of the eight blocks.

There is a total of $4(16)(17) + 4(16)(18) = 2,240$ experimental units. An analysis of variance for this situation is:

Source of variation	Degrees of freedom
Total	$[4(5 + 12) + 4(5 + 13)][4 + 12] = 2,240$
Correction for mean	1
Replicate = R	7
Maize genotypes = A	104
Maize control = MC	4
MC versus new = MCN	1
New = MN	99
$A \times R$	$4(7) = 28$
Bean genotypes	99
Bean control = BC	3
BC versus new = BCN	1
New = BN	95
$B \times R$	$3(7) = 21$
$A \times B$	1,896
$BC \times MC$	$3(4) = 12$
$BCN \times MC$	$1(4) = 4$
$BN \times MC$	$95(4) = 380$
$BC \times MCN$	$3(1) = 3$
$BN \times MCN$ within block	$11(1)(8) = 88$
$BCN \times MCN$ within block	$1(1)(8) = 8$
$BC \times MN$	$3(99) = 297$
$BCN \times MN$ within block	$4(1)(11) + 4(1)(12) = 92$
$BN \times MN$ within block	$4(11)(11) + 4(11)(12) = 1,012$
$A \times B \times R$	84

A NUMERICAL EXAMPLE

To illustrate the type of effects that are estimable from an augmented split block experiment design, a numerical example was constructed. The number of factor A treatments is $ac + an = 13 = 4$ controls + 9 new treatments. The number of factor B treatments is $bc + bn = 9 = 3$ controls + 6 new treatments. An augmented randomized complete block design with $r = 3$ blocks is used for both the factor A and factor B treatments. A systematic layout of the design and responses is given in Table 1.

An analysis of variance along with F-values is presented in Table 2. The sums of squares for the new treatments and the contrast of the new treatments versus the control treatments were pooled in the table. It is possible to obtain the sums of squares for each of the effects as given above but that was not done for the example.

Table 1. Systematic layout of an ASBED with artificial responses.

Block 1						Block2						Block3					
	B1	B2	B3	B4	B5		B1	B2	B3	B6	B7		B1	B2	B3	B8	B9
A1	9	7	5	9	1	A1	9	7	5	9	11	A1	11	10	10	15	9
A2	3	5	4	9	7	A2	13	5	4	9	7	A2	13	15	14	9	17
A3	8	7	5	9	2	A3	8	7	5	9	12	A3	8	17	15	9	12
A4	3	5	4	9	7	A4	13	5	4	9	7	A4	13	15	14	9	17
A5	8	7	5	9	2	A8	13	8	8	9	7	A11	13	15	14	19	17
A6	3	5	4	9	7	A9	8	7	8	9	12	A12	18	17	15	19	12
A7	8	7	5	9	2	A10	13	9	12	9	7	A13	13	15	14	19	17

Table2. Type III analysis of variance for the responses in Table 1.

Source of Variation	Degrees of freedom	Sum of squares	Mean square	F-value
Block = <i>R</i>	2	372.22	186.11	
Factor <i>A</i>	12	93.00	7.75	0.93
Control = <i>AC</i>		3	2.75	0.92
New = <i>AN</i>]	9	90.25	10.03
<i>AC</i> vs. <i>AN</i> = <i>ACN</i>				1.20
<i>A</i> × <i>R</i>	6	50.00	8.33	
<i>B</i>	8	123.79	15.47	0.71
Control = <i>BC</i>		2	21.56	10.78
New = <i>BN</i>]	6	102.23	17.04
<i>BC</i> vs. <i>BN</i> = <i>BCN</i>				0.78
<i>B</i> × <i>R</i>	4	87.44	21.86	
<i>A</i> × <i>B</i>	60	380.83	6.35	1.56
<i>AC</i> × <i>BC</i>	6	18.00	3.00	0.74
<i>AC</i> × <i>BN</i> + <i>AC</i> × <i>BCN</i>	18	150.80	8.38	2.05
<i>AN</i> × <i>BC</i> + <i>ACN</i> × <i>BC</i>	18	42.33	2.35	0.58
Other interactions	18	169.70	9.43	2.31
<i>AC</i> × <i>BC</i> × <i>R</i>	12	49.00	4.08	

To obtain the above analysis of variance table, four runs of SAS GLM were conducted. These were for the controls alone, the *A* controls and the *B* treatments, the *B* controls and the *A* treatments, and for the 105 observations. It is possible to obtain further partitioning of the sums of squares and degrees of freedom. The partitioning was described in the previous section.

Using the following SAS GLM model statement

Response = R A R*A B B*R A*B/solution,;

it is possible to obtain solutions for all effects that are present in the experiment. These are useful as the LSMEANS statement does not produce means for the new treatments. The effects obtained are presented in Table 3. No solutions are available for the empty spaces in the table. This is because there are combinations of new treatments that do not appear in a block. Solutions are possible only for combinations that appear in the design. When interpreting these values, the procedure used to obtain them must be understood. The SAS GLM procedure does not use the constraint that the sum of the effects for a factor is zero. Instead, the procedure uses the constraint that the highest number effect is set equal to zero. The consequence of this is that the highest numbered effect is subtracted from all the other effects. For interactions, the effects in the last row and in the last column are set to zero. That is, these effects are subtracted from all the others.

Table 3. SAS solutions for effects of the responses of Table 1.

<i>A</i> treatments	<i>B</i> treatments									<i>A</i> effect
	1	2	3	4	5	6	7	8	9	
1	4.44	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-8.00
2	-0.22	-0.05	0.28	2.25	7.25	7.58	7.58	-10.00	0.00	-0.00
3	2.00	5.83	5.17	3.92	3.92	12.58	17.58	-5.00	0.00	-5.00
4	-0.22	-0.06	0.28	2.25	7.25	7.58	7.58	-10.00	0.00	-0.00
5	0.75	0.50	-0.00	-0.00	0.00					2.25
6	-9.25	-6.50	-6.00	-5.00	0.00					7.25
7	0.75	0.50	0.00	0.00	0.00					2.25
8	-6.25	-5.50	-4.00			0.00	0.00			7.58
9	-16.25	-11.50	-9.00			-5.00	0.00			12.58
10	-6.25	-4.50	0.00			0.00	0.00			7.58
11	0.00	0.00	-0.00					0.00	0.00	-0.00
12	10.00	7.00	6.00					0.00	0.00	-5.00
13	0.00	0.00	0.00					0.00	0.00	0.00
<i>B</i> tr. Effect	-4.00	-2.00	-3.00	0.00	0.00	0.00	0.00	2.00	0.00	

The least squares means that appear on the SAS output are given in Table 4.

Table 4. Least squares means for the *A* and *B* control treatments.

<i>A</i> controls	<i>B</i> controls			Mean
	1	2	3	
1	9.67	8.00	6.67	8.11
2	9.67	8.33	7.33	8.44
3	8.00	10.33	8.33	8.89
4	9.67	8.33	7.33	8.44
Mean	9.25	8.75	7.42	8.47

The means for the new treatments are obtained as the response for a new treatment minus the block effect in which the new treatment occurred. The block effects may be obtained from an analysis of control data only. The new treatments do not contribute to estimating block effects as they appear only once in the experiment.

COMMENTS

Any experiment design may be augmented to accommodate a set of new treatments that are to be replicated once. The class of augmented experiment designs is very large and varied. Lack of material may be a factor in deciding to use only one replicate. Or, it may be that the experimenter has so many new lines that make it desirable to use only one replicate. For example, plant breeders of some crops may want to screen 30,000 new entries every year. Others screen 8,000 new entries each year. Producers of fungicides, herbicides, etc. may have hundreds of new entries for screening. Since some of these may kill all the plants in a plot, it is not desirable to do a screening on more than one experimental unit.

In field experiments involving intercrops, it is possible first to plant one crop in the field. Then, perpendicular to the planting of crop one, a second crop is planted across the first crop in each of the blocks, or possibly across the entire experiment. Such arrangements allow the experimenter to put the new treatments of two crops in each of the blocks in the desired manner to obtain two crop mixing or combining effects for the selected combinations. For three crops in a mixture, an experimenter could use one crop as the whole plots of a split plot design and then use an augmented split block design as above for the split plot treatments.

As stated, the use of an augmented split block experiment design allows the experimenter to screen new treatments for various cultural practices. The new treatments used in this situation would more than likely be new treatments that have survived previous stages of screening. Cultural practices such as soil preparation would be ideal for this design. Spraying fungicide or herbicide treatments would also fit into this design.

These designs are useful for any type of experiment that involves the screening of material where a number of conditions for screening are desired.

LITERATURE CITED

Federer, W. T. (1993). *Statistical Design and Analysis for Intercropping Experiments: Volume I: Two Crops*. Springer-Verlag, New York, Heidelberg, Berlin, Chapter 10.

Federer, W. T. (1999). *Statistical Design and Analysis for Intercropping Experiments: Volume II: Three or More Crops*. Springer-Verlag, New York, Heidelberg, Berlin.

Federer, W. T. (2002). Construction and analysis of an augmented lattice square design. *Biometrical Journal* 44(2):251-257.

Federer W, T. and B. T. Scully (1993). A parsimonious statistical design and breeding procedure for evaluating and selecting desirable characteristics over environments. *Theoretical and Applied Genetics* 86:612-620.

Mejza, Iwona (1998). Characterization of certain split-block designs with a control. *Biometrical Journal* 40(5):627-639.